

Carboxy terminus  
of  $\alpha$ -factor:

Gly -- Gln -- Pro -- Met -- Tyr COOH

Possible codons  
and their usage

GGU (90)	CAA (20)	CCA (32)	AUG (20)	UAC (33)
GGC (3)	CAG (0)	CCU (3)		UAU (0)
GGA (0)		CCC (1)		
GGG (0)		CCG (0)		

Consensus  
oligonucleotides:

5'-GG<sup>T</sup><sub>C</sub>CAACC<sup>A</sup><sub>T</sub>TATGTAC

Synthesized  
oligonucleotide  
pools complemen-  
tary to above:

I. 5'-GTACATTGGTTG<sup>A</sup><sub>G</sub>CC  
II. 5'-GTACATAGGTTG<sup>A</sup><sub>G</sub>CC

FIGURE 1

10/148946

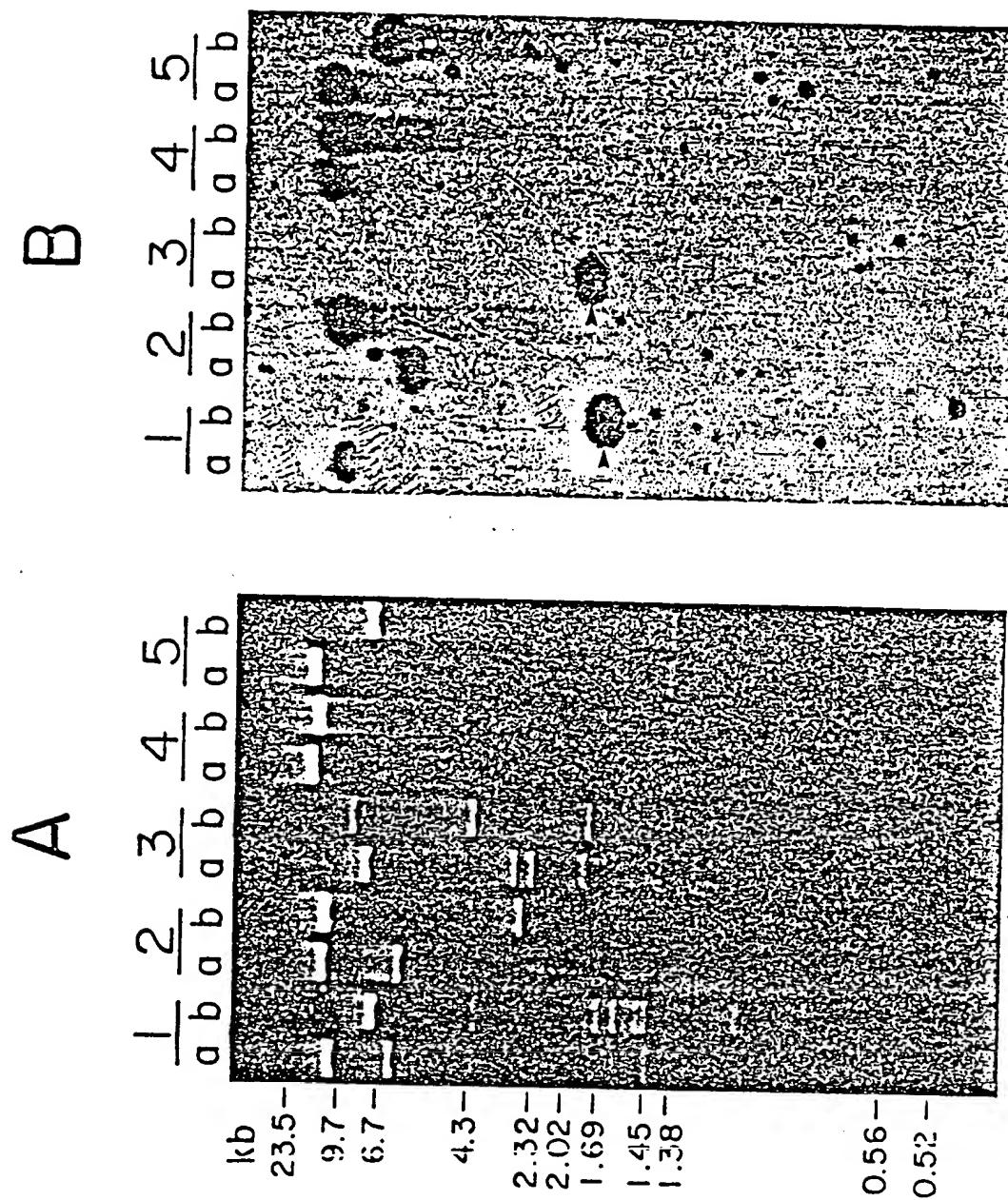


FIGURE 2

CGACAGTAAATTGCCGATTTAATAGCTCTACTGAAAAACAGTGGACCATGTGAAAAGATGCATCTCATTATCAA  
 -280 -260 -240 -220

ACACATAATTCAGTGAGCCCTTACTTCATTGATTGAAGTGCAAGAAAACCAAAAGCAACAACAGGTTGGATA  
 -200 -180 -160 -140

AGTACATATATAAGGGCCTTGTCCATCAAAATGTTACTGTTCTACGATTACGATTCAAGAATAGTT  
 -120 -100 -80 -60

1  
 Met Arg Phe Pro Ser Ile  
 CAAACAAGAAGATTACAAACTATCAATTCTACACAAATAAACGATTAAGA ATG AGA TTT CCT TCA ATT  
 -40 -20 1

20  
 Phe Thr Ala Val Leu Phe Ala Ala Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr  
 TTT ACT GCA GTT TTA TTC GCA GCA TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA  
 20 40 60

40  
 Glu Asp Glu Thr Ala Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Leu Asp Leu Glu Gly  
 GAA GAT GAA ACG GCA CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TTA GAT TTA GAA GGG  
 80 100 :120

60  
 Asp Phe Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe Ile  
 GAT TTC GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG TTT ATA  
 140 160 180

80  
 Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu Asp Lys Arg Glu  
 AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT TTG GAT AAA AGA GAG  
 200 220 240

100  
 Ala Glu Ala Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr Lys Arg Glu Ala  
 GCT GAA GCT TGG CAT TGG CAA CTA AAA CCT GGC CAA CCA ATG TAC AAG AGA GAA GCC  
 260 280 300

120  
 Glu Ala Glu Ala Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr Lys Arg Glu  
 GAA GCT GAA GCT TGG CAT TGG CTG CAA CTA AAG CCT GGC CAA CCA ATG TAC AAA AGA GAA  
 320 340 360

140  
 Ala Asp Ala Glu Ala Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr Lys Arg  
 GCC GAC GCT GAA GCT TGG CAT TGG CTG CAA CTA AAG CCT GGC CAA CCA ATG TAC AAA AGA  
 380 400 420

160 165  
 Glu Ala Asp Ala Glu Ala Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr End  
 GAA GCC GAC GCT GAA GCT TGG CAT TGG TTG CAG TTA AAA CCC GGC CAA CCA ATG TAC TAA  
 440 460 480

GCGCGACTGATAACAACAGTGTAGATGTAACAAAGTCGACTTTGTTCCACTGTACTTTAGCTCGTACAAAATACAAT  
 500 520 540 560

ATACTTTCTTCTCCGAAACAACTGTTTCCATGTAATATCCTTTCTATTTTCTCGTACCGTTACAAACTTAC  
 580 600 620 640

ACATACTTATATAGCTATTCACTTCTATACACTAAAAAAACTAAGACAATTAACTTATTTGCTGCGCTGCCATATTCAT  
 660 680 700 720

TTGTTATAATTCTATAATTCTATTAGTAGCTAAAAAAAGATGAATGTGAATCGAATCCTAAGAGAATT  
 740 760 780 800

FIGURE 3

TTCTCATTGGTACATCAATGCCAGCAACGATGCGCATCTGGCGACGCCTGAGTGATTGTTCAAGGTATCGAG  
 -300 -280 -260 -240

CCAAACATTATCGTTACTGTTCAAATATTCAAGTGTGTTCACTACAGAGTCGCCGTGGACCTAGTGAAACTTGGTGT  
 -220 -200 -180 -160

CTTACAGCGCAGAGACGAGGGCTTATATGTATAAAAGCTGCTTGATTCTGGTGTAGTTGAGGTGTCCTTCTATA  
 -140 -120 -100 -80

TCTGTTTATATCTATATAATGGATAATTACCATCACCTGCATCAAATTCCAGTAAAATTACATATTGGAGAAA  
 -60 -40 -20

1	10	20
Met Lys Phe Ile Ser Thr Phe Leu Thr Phe Ile Leu Ala Ala Val Ser Val Thr Ala Ser		
ATG AAA TTC ATT TCT ACC TTT CTC ACT TTT ATT TTA GCG GCC GTT TCT GTC ACT GCT AGT		
1 20 40 60		

30	40
Ser Asp Glu Asp Ile Ala Gln Val Pro Ala Glu Ala Ile Ile Gly Tyr Leu Asp Phe Gly	
TCC GAT GAA GAT ATC GCT CAG GTG CCA GCC GAG GCC ATT ATT GGA TAC TTG GAT TTC GGA	
80 100 120	

50	60
Gly Asp His Asp Ile Ala Phe Leu Pro Phe Ser Asn Ala Thr Ala Ser Gly Leu Leu Phe	
GGT GAT CAT GAC ATA GCT TTT TTA CCA TTC AGT AAC GCT ACC GCC AGT GGG CTA TTG TTT	
140 160 180	

70	80
Ile Asn Thr Thr Ile Ala Glu Ala Ala Glu Lys Glu Gln Asn Thr Thr Leu Ala Lys Arg	
ATC AAC ACC ACT ATT GCT GAG GCG GCT GAA AAA GAG CAA AAC ACC ACT TTG GCG AAA AGA	
200 220 240	

90	100
Glu Ala Val Ala Asp Ala Trp His Trp Leu Asn Leu Arg Pro Gly Gln Pro Met Tyr Lys	
GAG GCT GTT GCC GAC GCT TGG CAC TGG TAA ATT TTG AGA CCA GGC CAA CCA ATG TAC AAG	
260 280 300	

110	120
Arg Glu Ala Asn Ala Asp Ala Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr	
AGA GAG GCC AAC GCT GAT GCT TGG CAC TGG TTG CAA CTC AAG CCA GGC CAA CCA ATG TAC	
320 340 360	

End

TGA AAAATGACCCTAAACTACTTCTAAACCTCTCGATTCTTACGTTCATACAACACCTAGTTTATTTC  
 380 400 420

TTTCAATCTGAGTAGTTGAGTTTCGATCACTCACATAGAACATTTTTGCCATTAAATAAGTATTCTCTCAAAT  
 440 460 480 500

GATGCGATACTATAATACTCTTGCCATATATTACATTCACTAAATAGGCATGTTCTATATCCGTTCCGATTC  
 520 540 560 580

TGCTGCAAGCAAGGTTCCCTATCATTACCGGATTGTTCACTATGGTTGGAGCTC  
 600 620 640

FIGURE 4

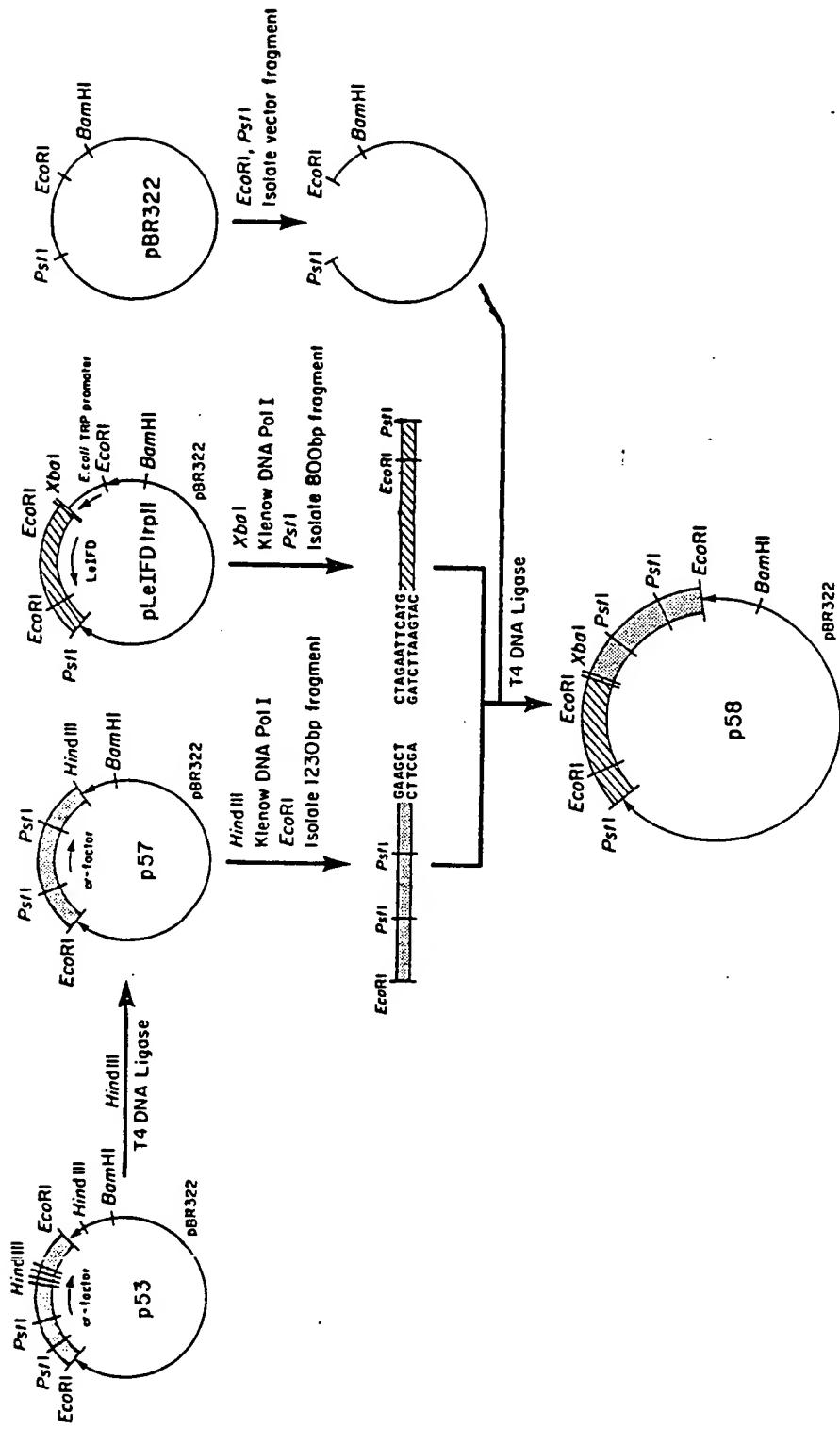


FIGURE 5

88-148946

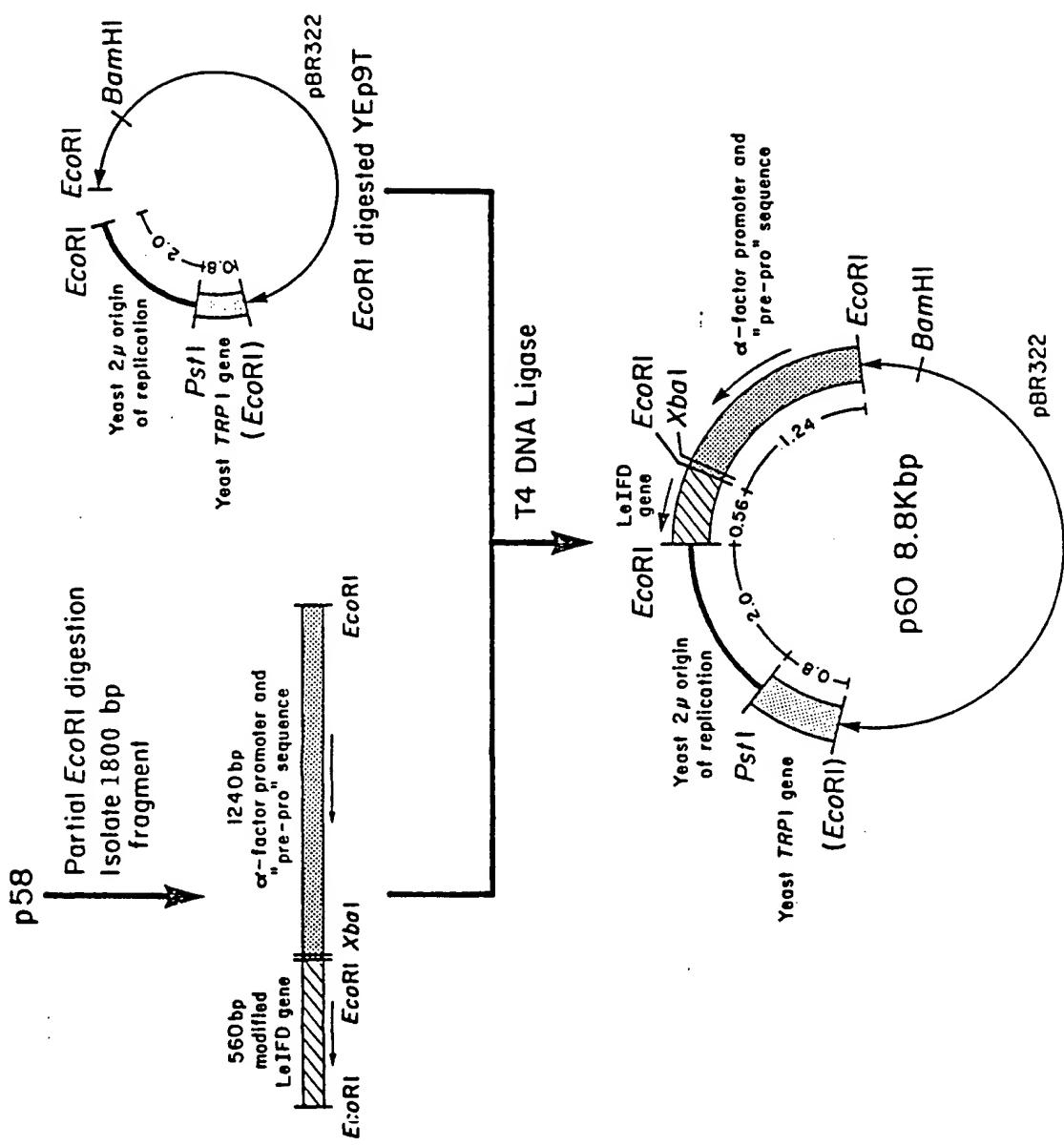


FIGURE 6

ALPHA-FACTOR "PRE-PRO" SEQUENCE                    MODIFIED IFN-ALPHA<sub>1</sub> GENE

...LYS	ARG	GLU	ALA	GLU	ALA	LEU	GLU	PHE	MET...
...AAA	AGA	GAG	GCT	GAA	GCT	CTA	GAA	TTC	ATG...
...TTT	TCT	CTC	CGA	CTT	CGA	GAT	CTT	AAG	TAC...

XbaI      EcoRI

FIGURE 7

M 148946

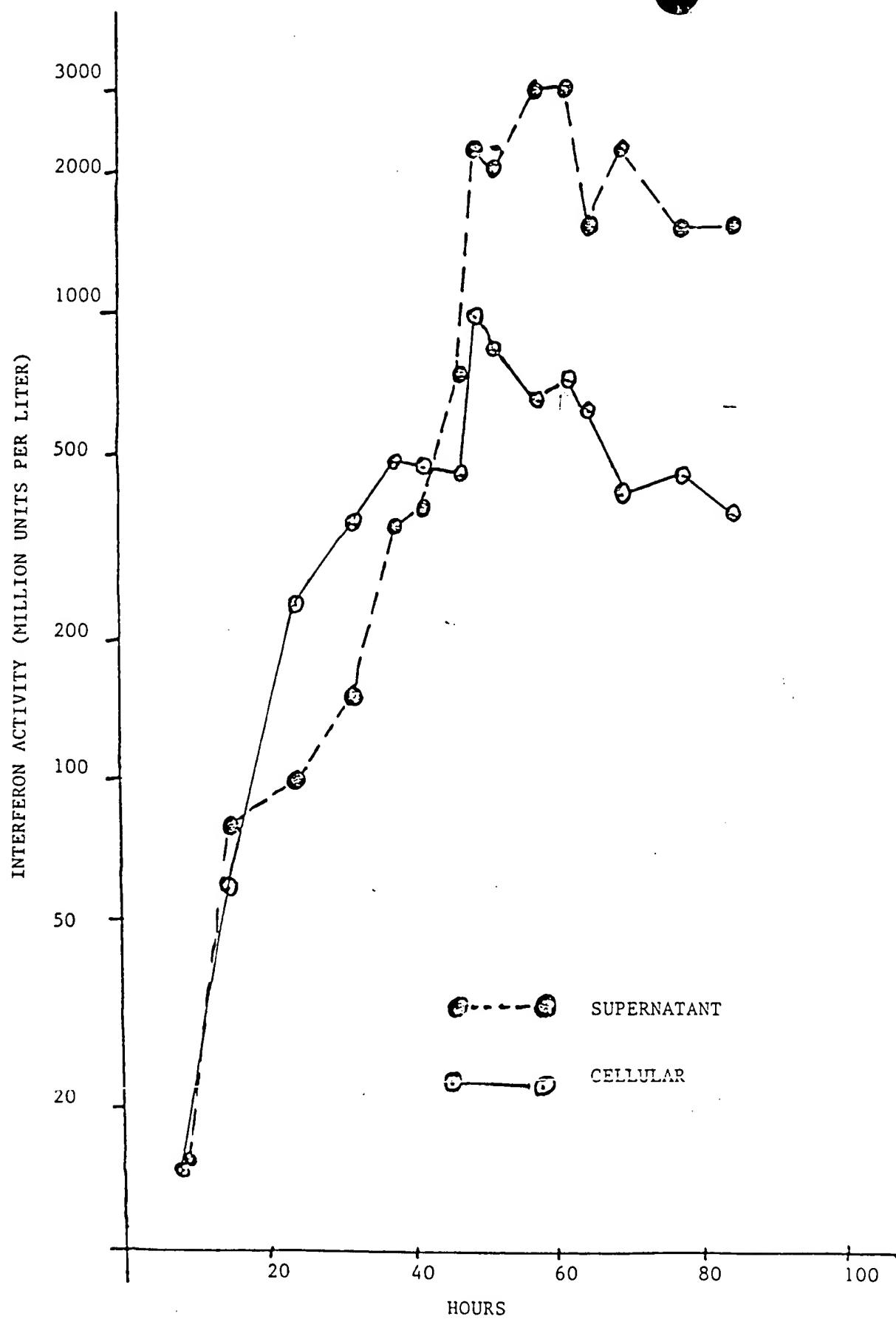


FIGURE 8

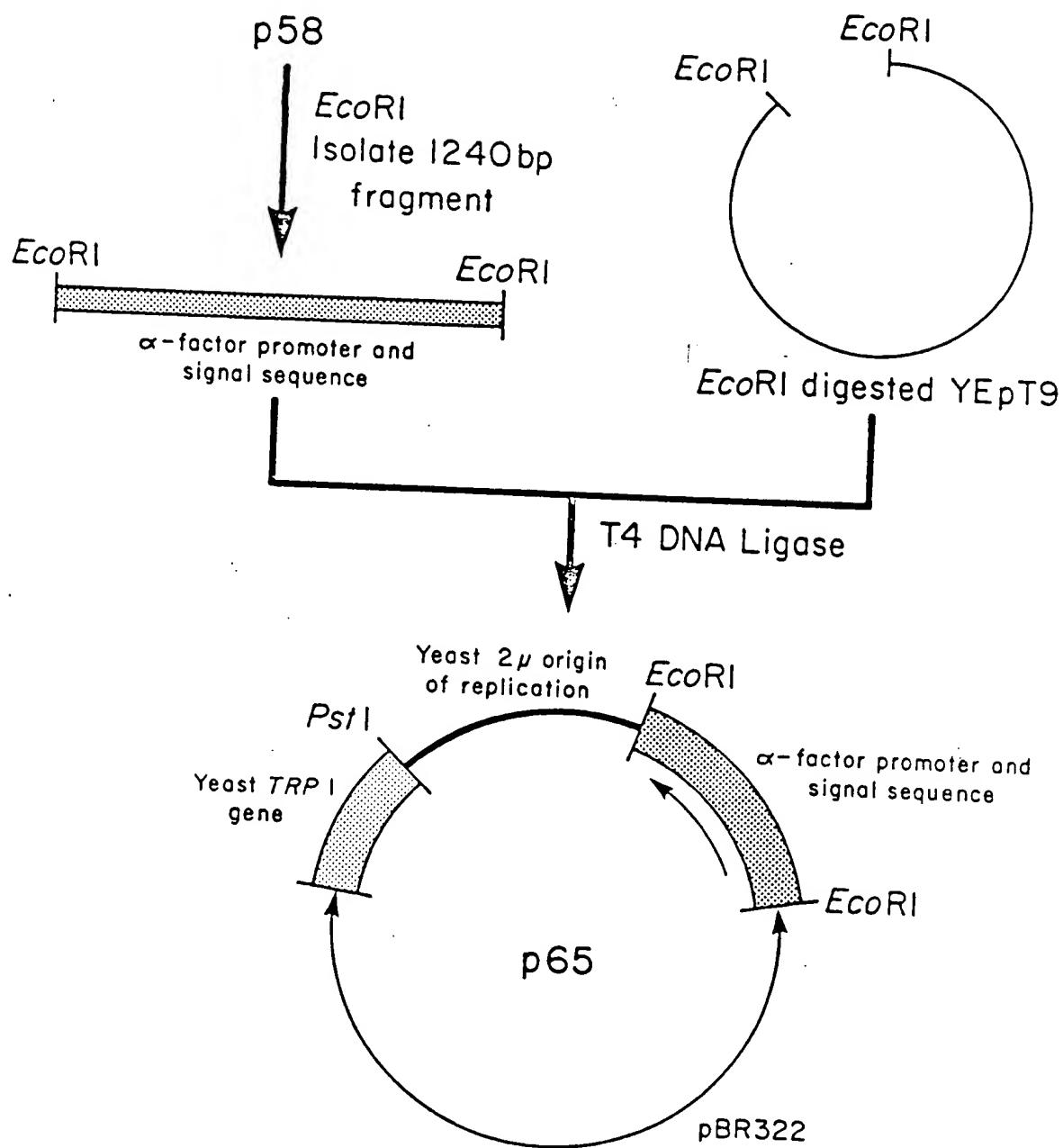


FIGURE 9

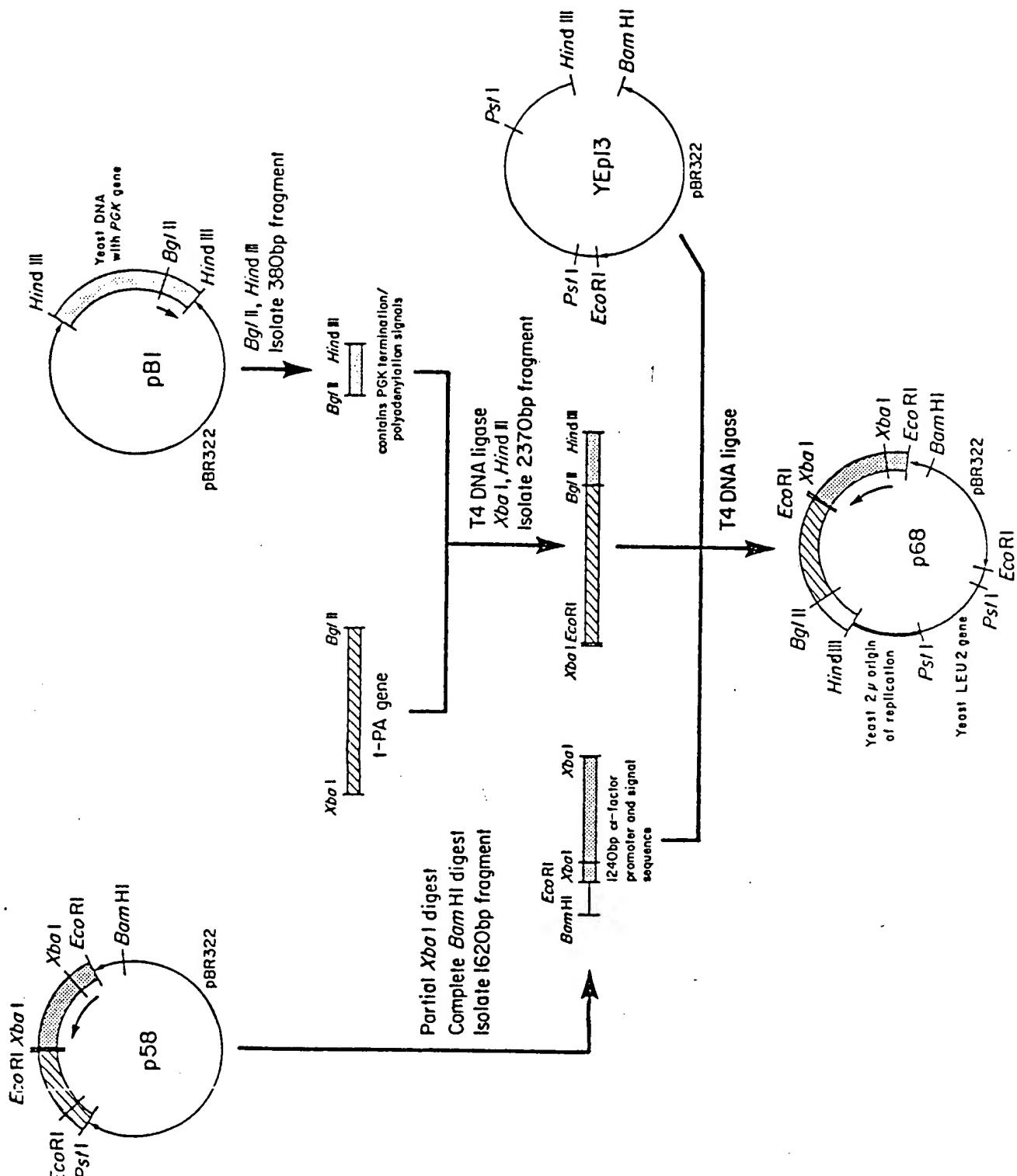


FIGURE 10

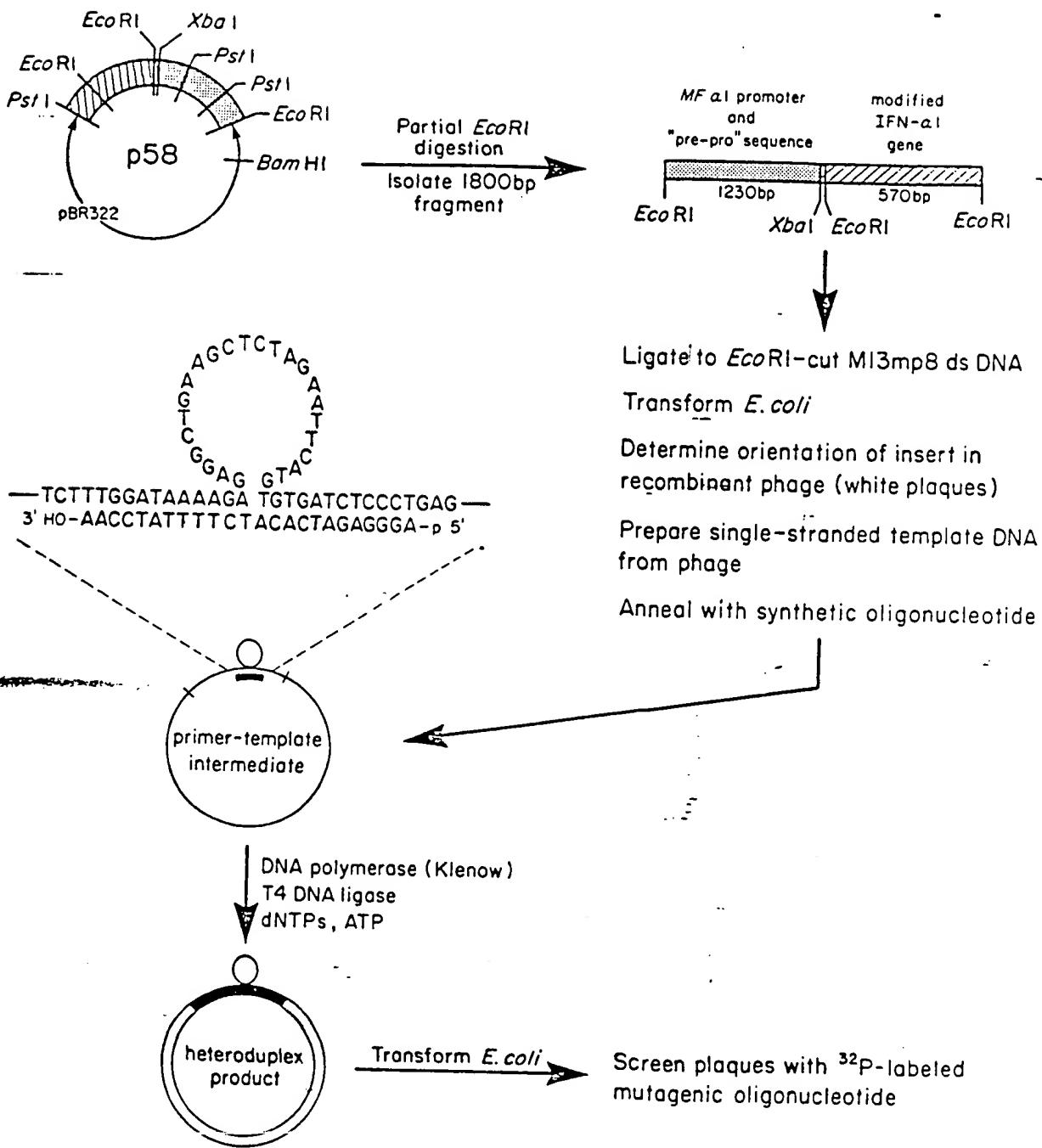


FIGURE 11

	1	10	20
<u>MF<math>\alpha</math>1</u>	Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser Ala Leu Ala Ala		
<u>MF<math>\alpha</math>2</u>	Met Lys Phe Ile Ser Thr Phe Leu Thr Phe	-----	Ile Leu Ala Ala
<u>Consensus</u>	Met --- Phe --- Ser --- Phe	-----	Leu Ala Ala
		30	40
<u>MF<math>\alpha</math>1</u>	Pro Val Asn Thr --- Thr Thr Glu Asp Glu Thr Ala Gln Ile Pro Ala Glu Ala Val Ile		
<u>MF<math>\alpha</math>2</u>	Val Ser Val Thr Ala Ser Ser Asp Glu Asp Ile Ala Gln Val Pro Ala Glu Ala Ile Ile	-----	
<u>Consensus</u>	----- Thr ----- Ala Gln --- Pro Ala Glu Ala --- Ile		
		50	60
<u>MF<math>\alpha</math>1</u>	Gly Tyr Leu Asp Leu Glu Gly Asp Phe Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr		
<u>MF<math>\alpha</math>2</u>	Gly Tyr Leu Asp Phe Gly Gly Asp His Asp Ile Ala Phe Leu Pro Phe Ser Asn Ala Thr		
<u>Consensus</u>	Gly Tyr Leu Asp ----- Gly Asp --- Asp --- Ala --- Leu Pro Phe Ser Asn --- Thr		
		70	80
<u>MF<math>\alpha</math>1</u>	Asn Asn Gly Leu Leu Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly		
<u>MF<math>\alpha</math>2</u>	Ala Ser Gly Leu Leu Phe Ile Asn Thr Thr Ile Ala Glu Ala Ala Glu Lys Glu Gln Asn	-----	
<u>Consensus</u>	----- Gly Leu Leu Phe Ile Asn Thr Thr Ile Ala ----- Ala --- Lys Glu -----		
		90	100
<u>MF<math>\alpha</math>1</u>	Val Ser Leu Asp Lys Arg Glu Ala Glu ----- Ala Trp His Trp Leu Gln Leu Lys Pro		
<u>MF<math>\alpha</math>2</u>	Thr Thr Leu Ala Lys Arg Glu Ala Val Ala Asp Ala Trp His Trp Leu Asn Leu Arg Pro	-----	
<u>Consensus</u>	----- Leu --- Lys Arg Glu Ala ----- Ala Trp His Trp Leu --- Leu --- Pro		
		110	120
<u>MF<math>\alpha</math>1</u>	Gly Gln Pro Met Tyr Lys Arg Glu Ala Glu Ala Trp His Trp Leu Gln Leu Lys		
<u>MF<math>\alpha</math>2</u>	Gly Gln Pro Met Tyr Lys Arg Glu Ala Asn Ala Asp Ala Trp His Trp Leu Gln Leu Lys		
<u>Consensus</u>	Gly Gln Pro Met Tyr Lys Arg Glu Ala --- Ala --- Ala Trp His Trp Leu Gln Leu Lys		
		130	140
<u>MF<math>\alpha</math>1</u>	Pro Gly Gln Pro Met Tyr Lys Arg Glu Ala Asp Ala Glu Ala Trp His Trp Leu Gln Leu		
<u>MF<math>\alpha</math>2</u>	Pro Gly Gln Pro Met Tyr		
<u>Consensus</u>	Pro Gly Gln Pro Met Tyr -----		
		150	160
<u>MF<math>\alpha</math>1</u>	Lys Pro Gly Gln Pro Met Tyr Lys Arg Glu Ala Asp Ala Glu Ala Trp His Trp Leu Gln		
<u>Consensus</u>	-----		
		168	
<u>MF<math>\alpha</math>1</u>	Leu Lys Pro Gly Gln Pro Met Tyr		
<u>Consensus</u>	-----		

FIGURE 12